On large deviations in the Gaussian autoregressive process: stable, unstable and explosive cases

BERNARD BERCU

Laboratoire de Mathématiques, Équipe de Probabilités, Statistique et Modélisation, Bâtiment 425, Université de Paris-Sud, 91405 Orsay Cedex, France. E-mail: bernard.bercu@math.u-psud.fr

For the Gaussian autoregressive process, the asymptotic behaviour of the Yule–Walker estimator is totally different in the stable, unstable and explosive cases. We show that, irrespective of this trichotomy, this estimator shares quite similar large deviation properties in the three situations. However, in the explosive case, we obtain an unusual rate function with a discontinuity point at its minimum.

Keywords: autoregressive Gaussian process; estimation; large deviations

1. Introduction

Consider the autoregressive process

$$X_{n+1} = \theta X_n + \varepsilon_{n+1}, \tag{1.1}$$

where (ε_n) are independent and identically distributed as $N(0, \sigma^2)$. The process is said to be stable if $|\theta| < 1$, unstable if $|\theta| = 1$ and explosive if $|\theta| > 1$. The Yule–Walker estimator of the unknown parameter θ is given by

$$\tilde{\theta}_n = \frac{\displaystyle\sum_{k=1}^{n} X_k X_{k-1}}{\displaystyle\sum_{k=0}^{n} X_k^2}. \tag{1.2}$$

The asymptotic behaviour of the Yule–Walker estimator $\tilde{\theta}_n$ is completely different in the stable, unstable and explosive cases. A well-known differentiator is given by the so-called Fisher information or standardizing function (see White 1958, Section 3) which is n, n^2 and θ^{2n} in the stable, unstable and explosive cases, respectively. One might therefore conclude that the large deviation behaviour of $(\tilde{\theta}_n)$ is totally different in the three situations. In fact, the purpose of this paper is to show that, irrespective of this trichotomy, $(\tilde{\theta}_n)$ shares quite similar large deviation properties in the three cases. First of all, in the stable case, it is known from the important study of Mann and Wald (1943) that $\tilde{\theta}_n$ converges almost surely to θ and that

$\sqrt{n}(\tilde{\theta}_n - \theta) \Rightarrow N(0, 1 - \theta^2)$. The large deviation behaviour of $(\tilde{\theta}_n)$ was also recently established by Bercu *et al*. (1997; 2000).

Theorem 1.1. *Assume that* $|\theta| < 1$. *Moreover, assume that* X_0 *is independent of* (ε_n) *and distributed as* $N(0, \sigma^2/(1 - \theta^2))$. *Then, the sequence* $(\tilde{\theta}_n)$ *satisfies a large deviation principle* (LDP) *with speed n and good rate function*

$$I(c) = \begin{cases} \dfrac{1}{2} \log\left(\dfrac{1 + \theta^2 - 2\theta c}{1 - c^2}\right) & \text{if } c \in\ ]-1, 1[, \\ +\infty & \text{otherwise.} \end{cases} \tag{1.3}$$

More precisely, set

$$t_c = \frac{(\theta c - 1)(\theta - c)}{1 - c^2}, \qquad \sigma_c^2 = \frac{1 - c^2}{(1 + \theta^2 - 2\theta c)^2}. \tag{1.4}$$

Then we can find a sequence $(d_{c,k})$ *such that, for any integer* $p > 0$ *and n large enough, and for* $1 > c > \theta$,

$$P(\tilde{\theta}_n \geqslant c) = \frac{\exp(-nI(c))}{\sigma_c t_c \sqrt{2\pi n}} J(c)\left[1 + \sum_{k=1}^{p} \frac{d_{c,k}}{n^k} + \mathcal{O}\left(\frac{1}{n^{p+1}}\right)\right] \tag{1.5}$$

while, for $-1 < c < \theta$,

$$P(\tilde{\theta}_n \leqslant c) = \frac{-\exp(-nI(c))}{\sigma_c t_c \sqrt{2\pi n}} J(c)\left[1 + \sum_{k=1}^{p} \frac{d_{c,k}}{n^k} + \mathcal{O}\left(\frac{1}{n^{p+1}}\right)\right] \tag{1.6}$$

with

$$J(c) = \left(\frac{(1 - \theta^2)(1 + \theta^2 - 2\theta c)(1 - c^2)^2}{(\theta c - 1)^4}\right)^{1/2}.$$

Surprisingly, we shall establish results similar to (1.5) and (1.6) in explosive and unstable situations. Hereafter, we assume for the sake of simplicity that $X_0 = x$ with $x \in \mathbb{R}$. Moreover, as we shall only consider the distribution of $\tilde{\theta}_n$, we may take, without loss of generality, $\sigma^2 = 1$.

2. Main results

First of all, we shall focus our attention on the more attractive explosive case $|\theta| > 1$. Relation (1.1) can be rewritten as

$$X_n = \theta^n x + \sum_{k=1}^{n} \theta^{n-k}\varepsilon_k. \tag{2.1}$$

It follows from (2.1) that, if $Y_n = \theta^{-n}X_n$, (Y_n) converges almost surely and in mean square to

$$Y = x + \sum_{k=1}^{\infty} \theta^{-k} \varepsilon_k. \tag{2.2}$$

Thus, we directly obtain via Toeplitz's lemma that almost surely

$$\frac{1}{\theta^{2n}} \sum_{k=0}^{n} X_k^2 \to \frac{\theta^2}{\theta^2 - 1} Y^2, \tag{2.3}$$

which immediately implies that $\tilde{\theta}_n$ converges almost surely to $1/\theta$. It is also known from the important paper of White (1958) and its extensions (Anderson 1959; Rao 1961; Touati 1996) that the asymptotic distribution of $(\tilde{\theta}_n)$, properly normalized, can be explicitly calculated (see White 1958, (4.25)). In the particular case $x = 0$, it coincides with the Cauchy distribution. Therefore, in view of all the above results, one can expect that the explosive case needs quite a different strategy than the stable case. In fact, we shall establish a sharp large deviation result for $(\tilde{\theta}_n)$ via the same approach as developed in Bercu *et al.* (2000) for the stable situation. It was inspired by the original work of Bahadur and Rao (1960) for the empirical mean of independent and identically distributed random variables. The large deviation properties of $(\tilde{\theta}_n)$ are related to those of

$$Z_n(c) = \sum_{k=1}^{n} X_k X_{k-1} - c \sum_{k=0}^{n} X_k^2$$

with $c \in \mathbb{R}$, since $P(\tilde{\theta}_n \geqslant c) = P(Z_n(c) \geqslant 0)$. If the LDP for $(Z_n(c))$ can be obtained for all $c \in \mathbb{R}$, the LDP for $(\tilde{\theta}_n)$ will immediately follow. Observe that the threshold c for $\tilde{\theta}_n$ acts as a parameter for $Z_n(c)$. From the Cauchy–Schwarz inequality, as $|\tilde{\theta}_n| < 1$, we shall only focus our attention on the interval $c \in]-1, 1[$. We can rewrite $Z_n(c)$ as $Z_n(c) = X_1 x - cx^2 + X^{(n)\mathrm{T}} M_n X^{(n)}$ where $X^{(n)\mathrm{T}} = (X_1, \ldots, X_n)$ and M_n is the tridiagonal Toeplitz matrix

$$M_n = \frac{1}{2} \begin{pmatrix} -2c & 1 & 0 & \ldots \\ 1 & -2c & 1 & \ldots \\ \ldots & \ldots & \ldots & \ldots \\ \ldots & 1 & -2c & 1 \\ \ldots & 0 & 1 & -2c \end{pmatrix}. \tag{2.4}$$

A classical tool for proving the LDP for $(Z_n(c))$ is the moment generating function

$$L_n(t) = \frac{1}{n} \log \mathrm{E}[\exp(t Z_n(c))], \tag{2.5}$$

where the parameter c is omitted in order to simplify the notation.

Lemma 2.1. *Let $\mathscr{D}$ be the effective domain of the limit L of L_n. Set*

$$c_1 = \frac{1 + \theta}{2\theta}, \qquad c_2 = \frac{1 - \theta}{2\theta},$$

$$t_1 = \frac{(1 - \theta)^2}{2(1 - c)}, \qquad t_2 = -\frac{(1 + \theta)^2}{2(1 + c)}, \qquad t_3 = 2\theta(\theta c - 1). \tag{2.6}$$

We have $\mathcal{D} = [0, t_1]$ if $c_1 \leqslant c < 1$, while $\mathcal{D} = [0, t_3]$ if $1/\theta \leqslant c \leqslant c_1$. Moreover, we have $\mathcal{D} = [t_3, 0]$ if $c_2 \leqslant c \leqslant 1/\theta$, while $\mathcal{D} = [t_2, 0]$ if $-1 < c \leqslant c_2$.

Remark. One can observe that the origin never belongs to the interior of $\mathcal{D}$. Consequently, a direct application of the Gärtner–Ellis theorem (see, for example, Dembo and Zeitouni 1998, Theorem 2.3.6) is not possible for proving an LDP for $(\tilde{\theta}_n)$.

For all $t \in \mathcal{D}$, set $p(t) = 1 + \theta^2 + 2ct$, $q(t) = -\theta - t$ and denote by $\delta(t)$ the square root of $p^2(t) - 4q^2(t)$. In addition, set

$$a(t) = \frac{p(t) + \delta(t)}{2}, \qquad b(t) = \frac{p(t) - \delta(t)}{2}. \tag{2.7}$$

Lemma 2.2. *For any t in the interior of $\mathcal{D}$, we have*

$$L_n(t) = L(t) + \frac{1}{n} H(t) + \frac{1}{n} R_n(t), \tag{2.8}$$

where

$$L(t) = -\frac{1}{2} \log a(t), \tag{2.9}$$

$$H(t) = -\frac{1}{2} \log\left(\frac{a(t) - \theta^2}{a(t) - b(t)}\right) - \frac{x^2}{2}(a(t) - 1), \tag{2.10}$$

$$R_n(t) = -\frac{1}{2} \log(1 - \eta(t)\rho(t)^n) - \frac{x^2}{2}\left(\frac{\delta(t)\eta(t)\rho(t)^n}{1 - \eta(t)\rho(t)^n}\right) \tag{2.11}$$

with $\rho(t) = b(t)/a(t)$ and $\eta(t) = (b(t) - \theta^2)/(a(t) - \theta^2)$. Moreover, as $0 \leqslant \rho(t) < 1$, the remainder R_n goes exponentially fast to zero, $R_n(t) = \mathcal{O}(\rho(t)^n)$.

Theorem 2.3. *The sequence $(\tilde{\theta}_n)$ satisfies an LDP with speed n and good rate function*

$$I(c) = \begin{cases} \frac{1}{2} \log\left(\dfrac{1 + \theta^2 - 2\theta c}{1 - c^2}\right) & \text{if } c \in \,]-1, 1[\text{ and } c \neq 1/\theta, \\ 0 & \text{if } c = 1/\theta, \\ +\infty & \text{otherwise.} \end{cases} \tag{2.12}$$

More precisely, we can find a sequence $(d_{c,k})$ such that, for any integer $p > 0$ and n large enough, and for $1 > c > 1/\theta$,

$$P(\tilde{\theta}_n \geqslant c) = \exp\left(-nI(c) - \frac{x^2(\theta - c)^2}{2(1 - c^2)}\right) \frac{J(c)}{\sigma_c t_c \sqrt{2\pi n}} \left[1 + \sum_{k=1}^{p} \frac{d_{c,k}}{n^k} + \mathcal{O}\left(\frac{1}{n^{p+1}}\right)\right] \tag{2.13}$$

while, for $-1 < c < 1/\theta$,

$$P(\tilde{\theta}_n \leqslant c) = -\exp\left(-nI(c) - \frac{x^2(\theta - c)^2}{2(1 - c^2)}\right)\frac{J(c)}{\sigma_c t_c \sqrt{2\pi n}}\left[1 + \sum_{k=1}^{p}\frac{d_{c,k}}{n^k} + \mathscr{O}\left(\frac{1}{n^{p+1}}\right)\right] \quad (2.14)$$

with t_c and σ_c^2 given by (1.4) and

$$J(c) = \left(\frac{(1 + \theta^2 - 2\theta c)(1 - c^2)}{(\theta c - 1)^2}\right)^{1/2}.$$

The coefficients $d_{c,1}$, $d_{c,2}$, ..., $d_{c,p}$ may be explicitly given as functions of the derivatives of L and H at point t_c. For example, the first coefficient $d_{c,1}$ is given by

$$d_{c,1} = \frac{1}{\sigma_c^2}\left(-\frac{h_2}{2} - \frac{h_1^2}{2} + \frac{l_4}{8\sigma_c^2} + \frac{l_3 h_1}{2\sigma_c^2} - \frac{5l_3^2}{24\sigma_c^4} + \frac{h_1}{t_c} - \frac{l_3}{2t_c\sigma_c^2} - \frac{1}{t_c^2}\right),$$

with $l_k = L^{(k)}(t_c)$ and $h_k = H^{(k)}(t_c)$.

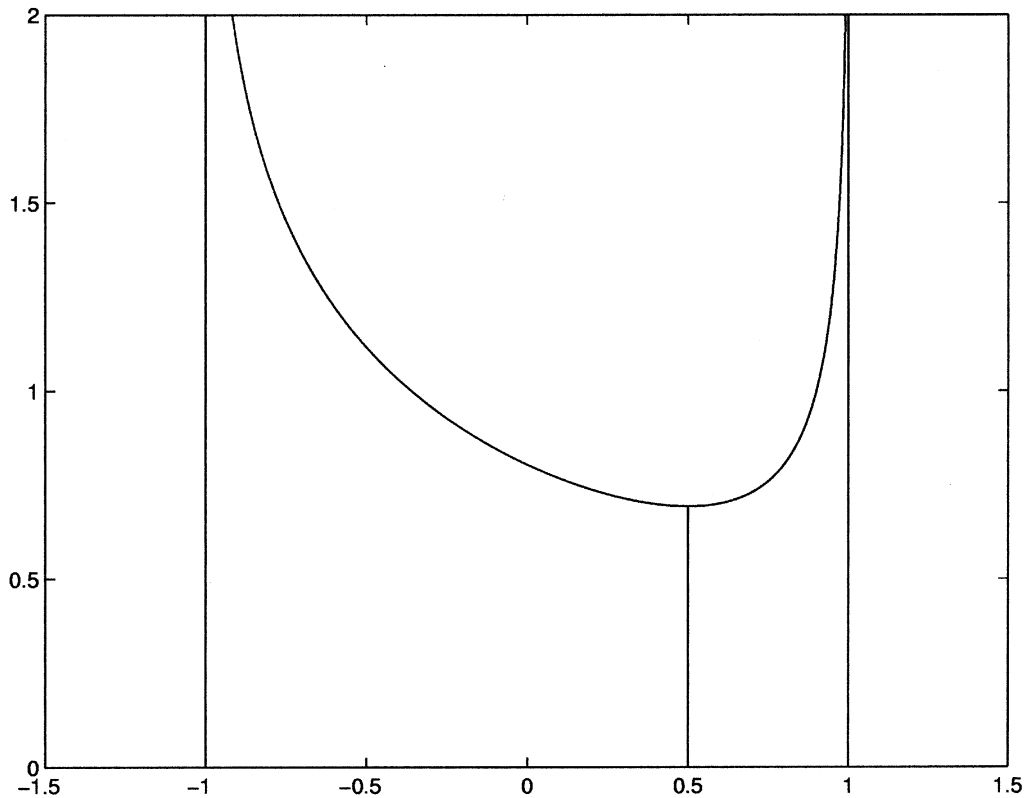Figure 1 shows the rate function I of (2.12) for $\theta = 2$.

Figure 1. The rate function I for $\theta = 2$

Remark. One can observe an unusual discontinuity point for the rate function I at its minimum $1/\theta$ due to the exponential rate of convergence $1/|\theta|^n$. Moreover, we have already seen that $\tilde{\theta}_n$ converges almost surely to $1/\theta$. Thus, we can rephrase Theorem 2.3 to say that the sequence $(\tilde{\theta}_n^{-1})$ satisfies an LDP with speed n and good rate function

$$K(c) = \begin{cases} \dfrac{1}{2} \log\left(\dfrac{c(c\theta^2 - 2\theta + c)}{c^2 - 1}\right) & \text{if } |c| > 1 \text{ and } c \neq \theta, \\ 0 & \text{if } c = \theta, \\ +\infty & \text{otherwise.} \end{cases}$$

Proof. We follow the same approach as recently given in Bercu *et al.* (2000) for quadratic forms of Gaussian stationary processes. However, observe that, as $|\theta| > 1$, the autoregressive process (1.1) is not stationary. We shall only establish (2.13) as the proof for (2.14) follows exactly the same lines as that for (2.13). First of all, it is easy to check from (2.9) that, for t in the interior of $\mathscr{D}$, $L'(t) = 0$ if and only if $t = t_c$, with

$$t_c = \frac{(\theta c - 1)(\theta - c)}{1 - c^2}. \tag{2.15}$$

One can also remark that, for all $1 > c > 1/\theta$, t_c always belongs to the interior of $\mathscr{D}$. Moreover, for all $1 > c > 1/\theta$, we have the decomposition $P(\tilde{\theta}_n \geqslant c) = A_n B_n$, with

$$A_n = \exp[nL_n(t_c)], \tag{2.16}$$

$$B_n = \mathrm{E}_n[\exp(-t_c Z_n)\mathbb{1}_{Z_n \geqslant 0}], \tag{2.17}$$

where E_n is the expectation over the usual change of probability

$$\frac{\mathrm{d}Q_n}{\mathrm{d}P} = \exp(t_c Z_n - nL_n(t_c)). \tag{2.18}$$

On the one hand, we have from (2.8) together with (2.16) that

$$A_n = \exp(nL(t_c) + H(t_c) + R_n(t_c)). \tag{2.19}$$

Furthermore, it follows from (2.10) and (2.11) respectively that

$$H(t_c) = -\frac{1}{2} \log\left(\frac{(\theta c - 1)^2}{(1 + \theta^2 - 2\theta c)(1 - c^2)}\right) - \frac{x^2}{2}\frac{(\theta - c)^2}{1 - c^2},$$

$$R_n(t_c) = -\frac{1}{2} \log(1 - \eta_c c^{2n}) - \frac{x^2}{2}\left(\frac{(1 + \theta^2 - 2\theta c)\eta_c c^{2n}}{1 - \eta_c c^{2n}}\right),$$

with $\eta_c = (c^2(1 + \theta^2 - 2\theta c) - \theta^2)/(\theta c - 1)^2$. Therefore, as $I(c) = -L(t_c)$, we deduce from (2.19) that

$$A_n = \exp\left(-nI(c) - \frac{x^2}{2}\frac{(\theta - c)^2}{1 - c^2}\right)J(c)[1 + \mathscr{O}(c^{2n})]. \tag{2.20}$$

On the other hand, we can rewrite (2.17) as

$$B_n = \mathrm{E}_n[\exp(-\sigma_c t_c \sqrt{n} U_n) \mathbb{1}_{U_n \geqslant 0}], \qquad \text{where } U_n = \frac{Z_n}{\sigma_c \sqrt{n}} \tag{2.21}$$

and $\sigma_c^2 = L''(t_c) = (1 - c^2)/(1 + \theta^2 - 2\theta c)^2$. Lemma 2.4 gives the asymptotic distribution of U_n and a Taylor expansion of B_n. Relation (2.13) follows from the conjunction of (2.20) and (2.22) which completes the proof of Theorem 2.3. $\square$

Lemma 2.4. *For all* $1 > c > 1/\theta$*, the distribution of* U_n *under* Q_n *converges, as n goes to infinity, to the standard* $N(0, 1)$ *distribution. Moreover, there exists a sequence* $(d_{c,k})$ *such that, for any integer* $p > 0$ *and n large enough,*

$$B_n = \frac{1}{\sigma_c t_c \sqrt{2\pi n}} \left[1 + \sum_{k=1}^{p} \frac{d_{c,k}}{n^k} + \mathcal{O}\left(\frac{1}{n^{p+1}}\right) \right]. \tag{2.22}$$

3. The unstable case

We shall now propose a sharp large deviation result for $(\tilde{\theta}_n)$ in the unstable case $|\theta| = 1$. First of all, it will be evident that the unstable case can be handled exactly as the explosive one since Lemma 2.2 remains valid in the unstable framework. The only important difference concerns the effective domain $\mathscr{D}$.

Lemma 3.1. *Let* $\mathscr{D}$ *be the effective domain of the function L. For* $\theta = 1$*, we have* $\mathscr{D} = [t_3, 0]$ *if* $0 \leqslant c < 1$*, while* $\mathscr{D} = [t_2, 0]$ *if* $-1 < c \leqslant 0$*. Moreover, for* $\theta = -1$*, we have* $\mathscr{D} = [0, t_1]$ *if* $0 \leqslant c < 1$*, while* $\mathscr{D} = [0, t_3]$ *if* $-1 < c \leqslant 0$*.*

Theorem 3.2. *If* $\theta = 1$*, the sequence* $(\tilde{\theta}_n)$ *satisfies an LDP with speed n and good rate function*

$$I^+(c) = \begin{cases} \dfrac{1}{2} \log\left(\dfrac{2}{1 + c}\right) & \text{if } -1 < c \leqslant 1, \\ +\infty & \text{otherwise.} \end{cases} \tag{3.1}$$

More precisely, we can find a sequence $(f_{c,k})$ *such that, for any integer* $p > 0$ *and n large enough, and for* $-1 < c < 1$*,*

$$P(\tilde{\theta}_n \leqslant c) = -\exp\left(-nI^+(c) - \frac{x^2}{8\sigma_c^2}\right) \frac{J^+(c)}{\sigma_c t_c \sqrt{2\pi n}} \left[1 + \sum_{k=1}^{p} \frac{f_{c,k}}{n^k} + \mathcal{O}\left(\frac{1}{n^{p+1}}\right) \right] \tag{3.2}$$

with t_c *and* σ_c^2 *given by* (1.4) *and* $J^+(c) = \sqrt{2(1 + c)}$*. Moreover, if* $\theta = -1$*, the sequence* $(\tilde{\theta}_n)$ *satisfies an LDP with speed n and good rate function*

$$I^-(c) = \begin{cases} \dfrac{1}{2} \log\left(\dfrac{2}{1 - c}\right) & \text{if } -1 \leqslant c < 1 \\ +\infty & \text{otherwise.} \end{cases} \tag{3.3}$$

More precisely, we can find a sequence $(g_{c,k})$ such that, for any integer $p > 0$ and n large enough, and for $-1 < c < 1$,

$$P(\tilde{\theta}_n \geqslant c) = \exp\left(-nI^-(c) - \frac{x^2}{8\sigma_c^2}\right) \frac{J^-(c)}{\sigma_c t_c \sqrt{2\pi n}} \left[1 + \sum_{k=1}^{p} \frac{g_{c,k}}{n^k} + \mathcal{O}\left(\frac{1}{n^{p+1}}\right)\right] \qquad (3.4)$$

with t_c and σ_c^2 given by (1.4) and $J^-(c) = \sqrt{2(1-c)}$. The coefficients $(f_{c,k})$ and $(g_{c,k})$ may be explicitly calculated as in Theorem 2.3.

The rate functions described in Theorem 3.2 are shown in Figure 2: I^+ of (3.1) for $\theta = 1$ in Figure 2(a) and I^- of (3.3) for $\theta = -1$ in Figure 2(b).
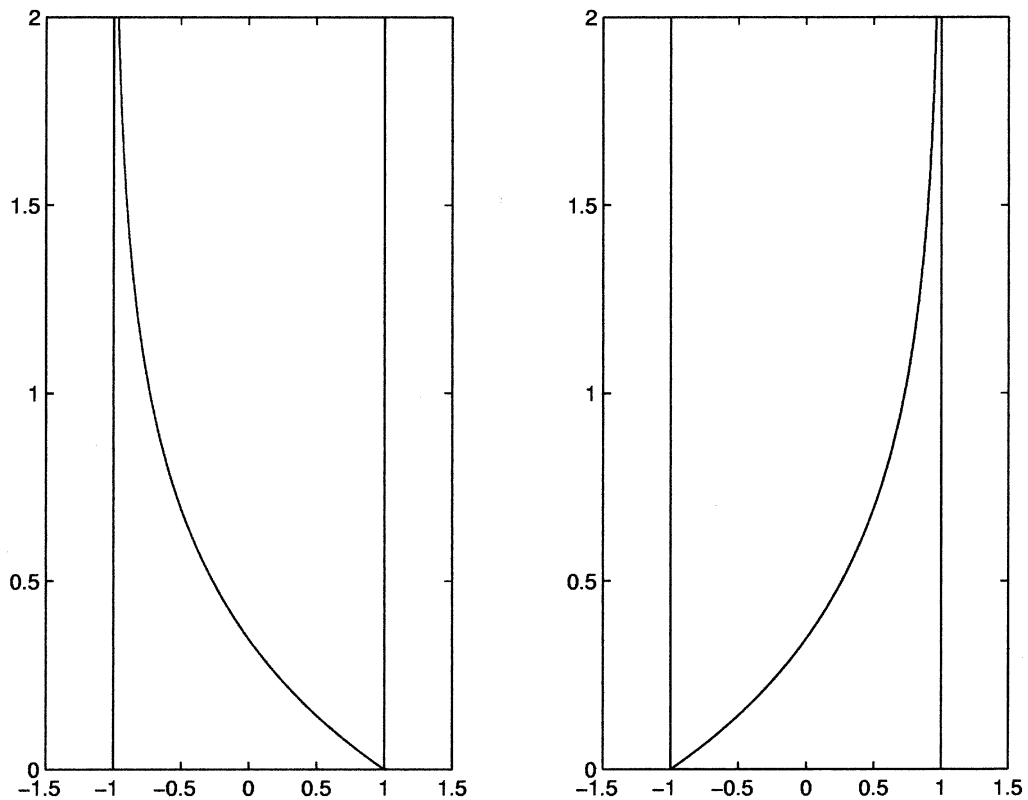

Figure 2. (a) The rate function I^+ for $\theta = 1$. (b) The rate function I^- for $\theta = -1$

4. A conditional approach

The Fisher information associated with (1.1) is n, n^2 and θ^{2n} in the stable, unstable and explosive cases, respectively. However, we have shown that, irrespective of this trichotomy,

$(\tilde{\theta}_n)$ shares similar large deviation properties with the same speed n in the three situations. Nevertheless, in the explosive case, it is also possible to find the expected speed θ^{2n} by proving an LDP for $(\tilde{\theta}_n)$ conditionally on the random variable Y given by (2.2). For all $(t, y) \in \mathbb{R}^2$, set

$$\Lambda_n(t, y) = \frac{1}{\theta^{2n}} \log \mathrm{E}[\exp(tZ_n(c))|Y = y], \tag{4.1}$$

where the parameter c is omitted in order to simplify the notation.

Lemma 4.1. *Let $\mathscr{S}$ be the effective domain of the limit Λ of Λ_n. Set*

$$d_1 = \frac{\theta + 1}{2}, \qquad d_2 = \frac{\theta - 1}{2},$$

$$t_1 = \frac{(1 - \theta)^2}{2(1 - c)}, \qquad t_2 = -\frac{(1 + \theta)^2}{2(1 + c)}, \qquad t_4 = 2(c - \theta). \tag{4.2}$$

If $1 < \theta < 3$, we have $\mathscr{S} = [t_4, t_1]$ if $d_2 \leqslant c < 1$, while $\mathscr{S} = [t_2, t_1]$ if $-1 < c \leqslant d_2$. Moreover, if $-3 < \theta < -1$, we have $\mathscr{S} = [t_2, t_1]$ if $d_1 \leqslant c < 1$, while $\mathscr{S} = [t_2, t_4]$ if $-1 < c \leqslant d_1$. Finally, if $|\theta| \geqslant 3$, we have $\mathscr{S} = [t_2, t_1]$.

Remark. One can observe that the effective domain $\mathscr{S}$ does not depend upon the variables x and y. Moreover, in contrast with the direct approach, it is also evident here that the origin always belongs to the interior of $\mathscr{S}$.

Lemma 4.2. *For any t in the interior of $\mathscr{S}$ and for all $y \in \mathbb{R}$, we have*

$$\Lambda_n(t, y) = y^2 \Lambda(t) + \frac{n}{\theta^{2n}} F(t) + \frac{1}{\theta^{2n}} G(t) + \frac{1}{\theta^{2n}} R_n(t) + P_n(t), \tag{4.3}$$

with $a(t)$ and $b(t)$ given by (2.7) and

$$\Lambda(t) = \frac{1}{2}(\theta^2 - 1)\left(\frac{\theta^2 - a(t)}{a(t) - 1}\right), \qquad F(t) = -\frac{1}{2} \log\left(\frac{a(t)}{\theta^2}\right), \tag{4.4}$$

$$G(t) = -\frac{1}{2} \log\left(\frac{a(t) - 1}{a(t) - b(t)}\right) + \frac{1}{2} \log((x - y)^2(\theta^2 - 1) - x^2(a(t) - 1)),$$

$$R_n(t) = -\frac{1}{2} \log(1 - \eta(t)\rho(t)^n),$$

$$P_n(t) = \frac{\pi(t)}{1 - \eta(t)\rho(t)^n}\left[\frac{y^2(\theta^2 - 1)^2\rho(t)^n}{a(t) - 1} - \frac{2xy(-1)^n v(t)^n}{\theta^n} + \frac{x^2(b(t) - 1)\rho(t)^n}{\theta^{2n}}\right],$$

where $\eta(t) = (b(t) - 1)/(a(t) - 1)$, $\pi(t) = (b(t) - a(t))/2(a(t) - 1)$ and $\rho(t) = b(t)/a(t)$, $v(t) = q(t)/a(t)$. Moreover, as $0 \leqslant \rho(t) < 1$ and $|v(t)| < 1$, the remainders P_n and R_n go exponentially fast to zero.

Theorem 4.3. *The sequence* $(\tilde{\theta}_n)$ *satisfies, conditionally on almost sure Y, an LDP with speed* θ^{2n} *and good rate function*

$$I_Y(c) = \begin{cases} \dfrac{1}{2} Y^2(\theta^2 - 1)\left(\dfrac{\theta c - 1}{\theta - c}\right)^2 & \text{if } c \in]-1, 1[, \\ +\infty & \text{otherwise.} \end{cases} \tag{4.5}$$

In particular, for $c \geq 1/\theta$, *we have almost surely*

$$\lim_{n \to \infty} \frac{1}{\theta^{2n}} \log P(\tilde{\theta}_n \geq c | Y) = -I_Y(c), \tag{4.6}$$

while, for $c \leq 1/\theta$, *we have almost surely*

$$\lim_{n \to \infty} \frac{1}{\theta^{2n}} \log P(\tilde{\theta}_n \leq c | Y) = -I_Y(c). \tag{4.7}$$

Proof. For all t in the interior of $\mathscr{S}$ and for all $y \in \mathbb{R}$, we have from (4.3)

$$\lim_{n \to \infty} \Lambda_n(t, y) = y^2 \Lambda(t) = \frac{1}{2} y^2 (\theta^2 - 1)\left(\frac{\theta^2 - a(t)}{a(t) - 1}\right).$$

Moreover, (2.7) implies that

$$\Lambda'(t) = -\frac{1}{2} y^2 \frac{(\theta^2 - 1)^2}{(a(t) - 1)^2}\left(c + \frac{cp(t) + 2q(t)}{\delta(t)}\right),$$

where $\delta(t)$ is the square root of $p^2(t) - 4q^2(t)$. Therefore, it is easy to see that the function Λ is steep as $|\Lambda'(t_1)| = |\Lambda'(t_2)| = |\Lambda'(t_4)| = +\infty$. In addition, for t in the interior of $\mathscr{S}$, $\Lambda'(t) = 0$ if and only if $t = t_c$, with t_c given by (1.4). Finally, Theorem 4.3 immediately follows from the Gärtner–Ellis theorem (see, for example, Dembo and Zeitouni 1998, Theorem 2.3.6) with $I_y(c) = -y^2 \Lambda(t_c)$. $\qquad\square$

Remark. It is also possible to find Theorem 2.3 by integrating Theorem 4.3 together with a careful analysis of the expansion (4.3). The important link between those two theorems is given by the functions L and F as, for all $t \in \mathscr{D}$, $F(t) = L(t) + \log|\theta|$.

5. Proofs

5.1. Proofs of Lemmas 2.1 and 2.2

It is easy to see that the joint distribution of $X^{(n)}$ is $N(m_n(x), \Gamma_n)$, where $m_n(x) = xv_n$,

$$\Gamma_n^{-1} = \begin{pmatrix} 1 + \theta^2 & -\theta & 0 & \dots \\ -\theta & 1 + \theta^2 & -\theta & \dots \\ \dots & \dots & \dots & \dots \\ \dots & -\theta & 1 + \theta^2 & -\theta \\ \dots & 0 & -\theta & 1 \end{pmatrix}, \qquad v_n = \begin{pmatrix} \theta \\ \theta^2 \\ \vdots \\ \theta^{n-1} \\ \theta^n \end{pmatrix}. \tag{5.1}$$

In order to evaluate the moment generating function (2.5), it is necessary to calculate the determinant of the tridiagonal matrix

$$D_n = \Gamma_n^{-1} - 2tM_n = \begin{pmatrix} p & q & 0 & \dots \\ q & p & q & \dots \\ \dots & \dots & \dots & \dots \\ \dots & q & p & q \\ \dots & 0 & q & r \end{pmatrix}, \tag{5.2}$$

with $p = 1 + \theta^2 + 2ct$, $q = -\theta - t$ and $r = p - \theta^2$, where the parameter t is omitted in order to simplify the notation. Denote by d_n the determinant of D_n. It immediately follows from (5.2) that $d_n = pd_{n-1} - q^2 d_{n-2}$ with $d_0 = 1$ and $d_1 = r$, so that

$$d_n = \begin{cases} \dfrac{(a - \theta^2)a^n - (b - \theta^2)b^n}{a - b} & \text{if } p^2 \neq 4q^2, \\[2ex] \left(\dfrac{p}{2}\right)^n \left(1 + \dfrac{(p - 2\theta^2)}{p} n\right) & \text{otherwise,} \end{cases} \tag{5.3}$$

with a and b given by (2.7). As in Bercu *et al.* (1997, Lemma 11) and Jensen (1995, Section 9.4), we obtain from (5.3) that, for any $t \in \mathscr{D}$, the matrix D_n is positive definite. One can also remark that, for $t \in \mathscr{D}$, $a = b$ so that $p^2 = 4q^2$ if and only if $t = t_1$ or $t = t_2$, and $a = \theta^2$ if and only if $t = 0$ or $t = t_3$, where t_1, t_2 and t_3 are given by (2.6). Therefore, for any t in the interior of $\mathscr{D}$, $a > b \geqslant 0$ and $a \neq \theta^2$.

After these algebraic preliminaries, we are now in position to prove Lemmas 2.1 and 2.2. Standard Gaussian calculations give that, for all $t \in \mathscr{D}$,

$$L_n(t) = -\frac{1}{2n} \log d_n - \frac{x^2}{2n} \left(p - 1 - q^2 \frac{d_{n-1}}{d_n}\right). \tag{5.4}$$

A similar expression can be found in White (1958, (4.20)). Moreover, it follows from (5.3) that, for any t in the interior of $\mathscr{D}$,

$$\log d_n = n \log a + \log\left(\frac{a - \theta^2}{a - b}\right) + \log(1 - \eta \rho^n), \tag{5.5}$$

$$\frac{d_{n-1}}{d_n} = \frac{1}{a} - \left(\frac{a - b}{ab}\right) \frac{\eta \rho^n}{1 - \eta \rho^n}, \tag{5.6}$$

with $\rho = b/a$ and $\eta = (b - \theta^2)/(a - \theta^2)$. Finally, (2.8) immediately follows from (5.4) together with (5.5) and (5.6), which completes the proof of Lemmas 2.1 and 2.2.

5.2. The alternative spectral expression

The goal of this subsection is to propose an alternative expression for the moment generating function (2.5) by rewriting Z_n as a weighted sum of independent and identically distributed random variables. Let $(\lambda_1^n, \ldots, \lambda_n^n)$ be the eigenvalues of $\Gamma_n^{1/2} M_n \Gamma_n^{1/2}$, with M_n and Γ_n given by (2.4) and (5.1) respectively, and set $\Lambda_n = \mathrm{diag}(\lambda_1^n, \ldots, \lambda_n^n)$. Denote by $(v_1^n, \ldots, v_n^n)$ an orthonormal basis of eigenvectors associated with $(\lambda_1^n, \ldots, \lambda_n^n)$ and let P_n be the orthogonal matrix with those vectors as column vectors. First of all, recall that $Z_n = X_1 x - c x^2 + X^{(n)\mathrm{T}} M_n X^{(n)}$. On the one hand, after an orthogonal change of basis, as the joint distribution of $X^{(n)}$ is $N(m_n(x), \Gamma_n)$, we can rewrite

$$X^{(n)\mathrm{T}} M_n X^{(n)} = z_n + \xi^{(n)\mathrm{T}} \Lambda_n \xi^{(n)} + 2\alpha_n(x)^{\mathrm{T}} \xi^{(n)}$$

where $z_n = m_n(x)^{\mathrm{T}} M_n m_n(x)$, $\alpha_n(x) = P_n^{\mathrm{T}} \Gamma_n^{1/2} M_n m_n(x)$ and where the distribution of $\xi^{(n)}$ is $N(0, I)$. On the other hand, we also have $X_1 = \theta x + \beta_n^{\mathrm{T}} \xi^{(n)}$ where $\beta_n = P_n^{\mathrm{T}} \Gamma_n^{1/2} u_n$ and $u_n^{\mathrm{T}} = (1, 0, \ldots, 0)$. Therefore, we obtain that

$$Z_n = z_n + x^2(\theta - c) + \xi^{(n)\mathrm{T}} \Lambda_n \xi^{(n)} + 2\gamma_n(x)^{\mathrm{T}} \xi^{(n)},$$

where $\gamma_n(x) = \alpha_n(x) + x\beta_n/2$. Finally, we find that

$$Z_n = z_n + x^2(\theta - c) + \sum_{i=1}^{n} \lambda_i^n (\xi_i^n)^2 + 2 \sum_{i=1}^{n} \gamma_i^n(x) \xi_i^n \tag{5.7}$$

where $\xi^{(n)\mathrm{T}} = (\xi_1^n, \ldots, \xi_n^n)$ and $\gamma_n(x)^{\mathrm{T}} = (\gamma_1^n(x), \ldots, \gamma_n^n(x))$.

Lemma 5.1. *The effective domain of L_n is*

$$\mathscr{D}_n = \{ t \in \mathbb{R} \,|\, \text{for all } 1 \leqslant i \leqslant n, \, 1 - 2t\lambda_i^n > 0 \}.$$

Moreover, for all $t \in \mathscr{D}_n$, we have

$$L_n(t) = \frac{t}{n}(z_n + x^2(\theta - c)) - \frac{1}{2n} \sum_{i=1}^{n} \log(1 - 2t\lambda_i^n) + \frac{2t^2}{n} \sum_{i=1}^{n} \frac{(\gamma_i^n(x))^2}{1 - 2t\lambda_i^n}, \tag{5.8}$$

and $L_n(t) = +\infty$ otherwise.

Remark. The proof of Lemma 5.1 immediately follows from (5.7) together with standard calculus on the $N(0, 1)$ distribution since $(\xi_1^n, \ldots, \xi_n^n)$ are independent and identically distributed as $N(0, 1)$. It is important to see that the eigenvalues $(\lambda_1^n, \ldots, \lambda_n^n)$ so that $\mathscr{D}_n$ do not depend upon the variable x.

5.2. The Szegö theorem

As $(\lambda_1^n, \ldots, \lambda_n^n)$ are the eigenvalues of $\Gamma_n^{1/2} M_n \Gamma_n^{1/2}$ and $\det \Gamma_n = 1$, we have

$$\sum_{i=1}^{n} \log(1 - 2t\lambda_i^n) = \log \det(\Gamma_n^{-1} - 2t M_n).$$

Moreover, $M_n = T_n(f)$, where $T_n(f)$ is the Toeplitz matrix associated with the function f defined, for all $x \in \mathbb{T} = [-\pi, \pi[$, by $f(x) = \cos x - c$. In addition, Γ_n^{-1} is quite similar to the Toeplitz matrix $T_n(g)$ with, for all $x \in \mathbb{T}$, $g(x) = \theta^{-2}(1 + \theta^2 - 2\theta \cos x)$. Denote by τ_n the empirical measure associated with $(\lambda_1^n, \ldots, \lambda_n^n)$,

$$\tau_n = \frac{1}{n} \sum_{i=1}^{n} \delta_{\lambda_i^n}.$$

Lemma 5.2. *For all $c \in\,] -1, 1[$, the sequence (τ_n) converges weakly to τ, which is the image probability of the uniform measure on the torus $\mathbb{T}$ by the applicaiton of $\varphi \in L^\infty(\mathbb{T})$ given, for all $x \in \mathbb{T}$, by*

$$\varphi(x) = \frac{\theta^2(\cos x - c)}{(1 + \theta^2 - 2\theta \cos x)}.$$

Proof. Lemma 5.2 immediately follows from the Szegö theorem on Toeplitz matrices (see Grenander and Szegö 1958, p. 65; and Bercu *et al.* 1997, Lemma 9) as for all $c \in\,] -1, 1[$ and for all $t \in \mathscr{D}$,

$$\lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^{n} \log(1 - 2t\lambda_i^n) = \frac{1}{2\pi} \int_{\mathbb{T}} \log(g(x) - 2tf(x))\, \mathrm{d}x$$

and by the Jacobi–Jensen formula,

$$\frac{1}{2\pi} \int_{\mathbb{T}} \log(1 + \theta^2 - 2\theta \cos x)\, \mathrm{d}x = \log \theta^2.$$

5.4. Proof of Lemma 2.4

Denote by Φ_n the characteristic function of U_n under the probability Q_n given by (2.18), $\Phi_n(u) = \mathrm{E}_n[\exp(\mathrm{i}uU_n)]$, so that

$$\Phi_n(u) = \exp\left[n\left(L_n\left(t_c + \frac{\mathrm{i}u}{\sigma_c \sqrt{n}}\right) - L_n(t_c)\right)\right]. \tag{5.9}$$

Then, we deduce from Lemma 5.1 that

$$|\Phi_n(u)|^2 = \prod_{i=1}^{n}\left(1 + \frac{4u^2(\lambda_i^n)^2}{\sigma_c^2 n(1 - 2t_c\lambda_i^n)^2}\right)^{-1/2} \exp\left(\frac{-4u^2\sigma_c^{-2}n^{-1}(\gamma_i^n(x))^2(1 - 2t_c\lambda_i^n)^{-1}}{(1 - 2t_c\lambda_i^n)^2 + 4u^2\sigma_c^{-2}n^{-1}(\lambda_i^n)^2}\right).$$

Therefore, since $1 - 2t\lambda_i^n > 0$, for all $1 \leq i \leq n$, we find that

$$|\Phi_n(u)|^2 \leq \prod_{i=1}^{n}\left(1 + \frac{4u^2(\lambda_i^n)^2}{\sigma_c^2 n(1 - 2t_c\lambda_i^n)^2}\right)^{-1/2}. \tag{5.10}$$

Hence, $\Phi_n \in L^2(\mathbb{R})$ as soon as $n \geqslant 2$. Then, we obtain by use of the Parseval formula in (2.21) that

$$B_n = \frac{1}{2\pi} \int_{\mathbb{R}} \left(\frac{1}{\sigma_c t_c \sqrt{n} + \mathrm{i}u} \right) \Phi_n(u)\, \mathrm{d}u = \frac{C_n}{\sigma_c t_c \sqrt{2\pi n}},$$

with

$$C_n = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \left(1 + \frac{\mathrm{i}u}{\sigma_c t_c \sqrt{n}} \right)^{-1} \Phi_n(u)\, \mathrm{d}u.$$

In order to prove (2.22), it remains to show that, for $1 > c > 1/\theta$, there exists a sequence $(d_{c,k})$ such that, for any integer $p > 0$ and n large enough,

$$C_n = 1 + \sum_{k=1}^{p} \frac{d_{c,k}}{n^k} + \mathcal{O}\left(\frac{1}{n^{p+1}} \right). \tag{5.11}$$

For some positive constant s, set $s_n = sn^{1/6}$. We split C_n into two terms, $C_n = I_n + J_n$, where

$$I_n = \frac{1}{\sqrt{2\pi}} \int_{|u| \leqslant s_n} \left(1 + \frac{\mathrm{i}u}{\sigma_c t_c \sqrt{n}} \right)^{-1} \Phi_n(u)\, \mathrm{d}u,$$

$$J_n = \frac{1}{\sqrt{2\pi}} \int_{|u| > s_n} \left(1 + \frac{\mathrm{i}u}{\sigma_c t_c \sqrt{n}} \right)^{-1} \Phi_n(u)\, \mathrm{d}u.$$

On the one hand, proceeding as in Bercu *et al.* (2000, Theorem 3.2), we obtain, via Lemma 5.2 together with (5.10), that

$$|J_n| = \mathcal{O}(\exp(-\mu n^{1/3})), \tag{5.12}$$

where μ is a positive constant. On the other hand, we find by the Taylor expansion of Φ_n established in Lemma 5.3 below that

$$I_n = \frac{1}{\sqrt{2\pi}} \int_{|u| \leqslant s_n} \exp\left(-\frac{u^2}{2} \right) \left[1 + \sum_{k=1}^{2p+1} \frac{\psi_k(u)}{(\sqrt{n})^k} \right] \mathrm{d}u + \mathcal{O}\left(\frac{1}{n^{p+1}} \right), \tag{5.13}$$

where ψ_k are polynomials in even powers of u for k even, and in odd powers of u for k odd. The initial sequence (φ_k) in (5.15) is replaced by (ψ_k) due to the fraction preceding Φ_n in I_n. Finally, we obtain (5.11) via standard calculus on the $N(0, 1)$ distribution.

5.5. A Taylor expansion

The goal of this subsection is to prove a Taylor expansion for Φ_n similar to that established for independent and identically distributed random variables by Cramér (1970, p. 72) and Esseen (1945, p. 44). First of all, it follows from Lemma 2.2 and (2.11) that for all $c \in\,] -1, 1[$ and for any $k \in \mathbb{N}$, $R_n^{(k)}(t_c) = \mathcal{O}(n^k c^{2n})$. Therefore, we obtain from (2.8) that

$$L_n^{(k)}(t_c) = l_k + \frac{1}{n} h_k + \mathcal{O}(n^{k-1} c^{2n}), \tag{5.14}$$

where $l_k = L^{(k)}(t_c)$ and $h_k = H^{(k)}(t_c)$. Then, via (5.9) together with (5.14), we obtain the following Taylor expansion for Φ_n.

Lemma 5.3. *For any $p \geqslant 0$ and for any $c \in \,]-1, 1[$, there exists a polynomial sequence (φ_k) independent of p, such that, for n large enough,*

$$\Phi_n(u) = \exp\left(-\frac{u^2}{2}\right)\left[1 + \sum_{k=1}^{2p+1} \frac{\varphi_k(u)}{(\sqrt{n})^k} + \mathcal{O}\left(\frac{\max(u^2, u^6)^{p+1}}{n^{p+1}}\right)\right], \tag{5.15}$$

where the remainder is uniform in u as soon as $|u| = \mathcal{O}(n^{1/6})$. Moreover, the φ_k are polynomials in even powers of u for k even and in odd powers of u for k odd. For example,

$$\varphi_1(u) = -\frac{\mathrm{i}u^3}{6\sigma_c^3}\,l_3 + \frac{\mathrm{i}u}{\sigma_c}\,h_1,$$

$$\varphi_2(u) = -\frac{u^2}{2\sigma_c^2}(h_2 + h_1^2) + \frac{u^4}{24\sigma_c^4}(l_4 + 4l_3 h_1) - \frac{u^6}{72\sigma_c^6}\,l_3^2.$$

Proof. We deduce from (5.9) and (5.14) that there exists $\xi \in \mathbb{R}$ such that, for any $c \in \,]-1, 1[$ and for all $|u| < \sqrt{n}$,

$$\log \Phi_n(u)$$

$$= \sum_{k=1}^{2p+3}\left(\frac{\mathrm{i}u}{\sigma_c\sqrt{n}}\right)^k\left(\frac{nl_k + h_k}{k!}\right) \tag{5.16}$$

$$+ \frac{n}{(2p+4)!}\left(\frac{u}{\sigma_c\sqrt{n}}\right)^{2p+4} L_n^{(2p+4)}(t_c + \mathrm{i}\xi) + \mathcal{O}(n^{2p+3}c^{2n}).$$

Moreover, it follows from (5.8) that, for all $t \in \mathscr{D}_n$ and for all $k \geqslant 3$,

$$L_n^{(k)}(t) = \frac{2^k k!}{2n}\left(\frac{1}{k}\sum_{i=1}^n \frac{(\lambda_i^n)^k}{(1 - 2t\lambda_i^n)^k} + \sum_{i=1}^n \frac{(\gamma_i^n(x))^2(\lambda_i^n)^{k-2}}{(1 - 2t\lambda_i^n)^{k+1}}\right). \tag{5.17}$$

Recall that $\gamma_n(x) = \alpha_n(x) + x\beta_n/2$, where $\alpha_n(x) = \Lambda w_n(x)$ with $w_n(x) = P_n^{\mathrm{T}}\Gamma_n^{-1/2}m_n(x)$ and $\beta_n = P_n^{\mathrm{T}}\Gamma_n^{1/2}u_n$. Set $\alpha_n(x)^{\mathrm{T}} = (\alpha_1^n(x), \ldots, \alpha_n^n(x))$ and $\beta_n^{\mathrm{T}} = (\beta_1^n, \ldots, \beta_n^n)$. On the one hand,

$$\sum_{i=1}^n\left(\frac{\alpha_i^n(x)}{\lambda_i^n}\right)^2 = \|w_n(x)\|^2 = \theta^2 x^2.$$

On the other hand,

$$\|\beta_n\|^2 = u_n^{\mathrm{T}}\Gamma_n u_n = 1.$$

Then, since $|L_n^{(2p+4)}(t_c + \mathrm{i}\xi)| \leqslant L_n^{(2p+4)}(t_c)$, for all $\xi \in \mathbb{R}$, we obtain from Lemma 5.2 together with (5.17) that

$$|L_n^{(2p+4)}(t_c + \mathrm{i}\xi)| = \mathcal{O}(1).$$

Therefore, as $l_1 = 0$ and $l_2 = \sigma_c^2$, we find, for all $|u| < \sqrt{n}$, that

$$\log \Phi_n(u) = -\frac{u^2}{2} + n \sum_{k=3}^{2p+3} \left(\frac{iu}{\sigma_c \sqrt{n}}\right)^k \frac{l_k}{k!} + \sum_{k=1}^{2p+1} \left(\frac{iu}{\sigma_c \sqrt{n}}\right)^k \frac{h_k}{k!} + \mathscr{O}\left(\frac{u^{2p+4}}{n^{p+1}}\right). \tag{5.18}$$

Finally, we obtain Lemma 5.3 by taking the exponential of both sides of (5.18), remarking that in the range $|u| = \mathscr{O}(n^{1/6})$ and for any $k \geqslant 3$, the quantity $nu^k/(\sqrt{n})^k$ remains bounded in (5.18). $\qquad \square$

5.6. Proofs of Lemmas 4.1 and 4.2

We shall follow essentially the same approach as that of Section 5.1 for the proofs of Lemmas 2.1 and 2.2. First of all, we can show after some linear algebra calculation that, for all $y \in \mathbb{R}$, the conditional distribution of $X^{(n)}$ given $Y = y$ is $N(\mu_n(x, y), \Delta_n)$, where $\mu_n(x, y) = m_n(x) + (y - x)\theta^{2n}(\theta^2 - 1)\Delta_n u_n$,

$$\Delta_n^{-1} = \begin{pmatrix} 1+\theta^2 & -\theta & 0 & \cdots \\ -\theta & 1+\theta^2 & -\theta & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & -\theta & 1+\theta^2 & -\theta \\ \cdots & 0 & -\theta & \theta^2 \end{pmatrix}, \qquad u_n = \begin{pmatrix} 0 \\ \vdots \\ \vdots \\ 0 \\ \theta^{-n} \end{pmatrix}. \tag{5.19}$$

The only difference between the two covariance matrices Γ_n and Δ_n is the last coefficient, which is 1 and θ^2 for Γ_n^{-1} and Δ_n^{-1}, respectively. This difference is crucial as $\det \Gamma_n = 1$ whereas $\det \Delta_n = \theta^{-2n}$. Let S_n and T_n be the two tridiagonal matrices

$$S_n = \begin{pmatrix} p & q & 0 & \cdots \\ q & p & q & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & q & p & q \\ \cdots & 0 & q & r \end{pmatrix}, \qquad T_n = \begin{pmatrix} p & q & 0 & \cdots \\ q & p & q & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & q & p & q \\ \cdots & 0 & q & p \end{pmatrix}, \tag{5.20}$$

with $p = 1 + \theta^2 + 2ct$, $q = -\theta - t$ and $r = p - 1$. If s_n and t_n denote the determinants of S_n and T_n, we find exactly as in Section 5.1 that

$$t_n = \begin{cases} \dfrac{a^{n+1} - b^{n+1}}{a - b} & \text{if } p^2 \neq 4q^2, \\ \left(\dfrac{p}{2}\right)^n (1 + n) & \text{otherwise,} \end{cases} \tag{5.21}$$

$$s_n = \begin{cases} \dfrac{(a - 1)a^n - (b - 1)b^n}{a - b} & \text{if } p^2 \neq 4q^2, \\ \left(\dfrac{p}{2}\right)^n \left(1 + \dfrac{(p - 2)}{p} n\right) & \text{otherwise,} \end{cases} \tag{5.22}$$

with a and b given by (2.7). Moreover, we obtain from (5.22) that, for any $t \in \mathscr{S}$, the matrix

S_n is positive definite. One can also remark that for $t \in \mathscr{S}$, $a = b$ so that $p^2 = 4q^2$ if and only if $t = t_1$ or $t = t_2$, and $a = 1$ if and only if $t = t_4$, where t_1, t_2 and t_4 are given by (4.2). Therefore, for any t in the interior of $\mathscr{S}$, $a > b \geqslant 0$ and $a \neq 1$. Furthermore, we find after some standard Gaussian calculations that, for any t in the interior of $\mathscr{S}$,

$$\Lambda_n(t, y) = \frac{1}{2\theta^{2n}}(n \log \theta^2 - \log s_n) \tag{5.23}$$

$$- \frac{x^2}{2\theta^{2n}}(p - 1) - \frac{y^2}{2}(\theta^2 - 1) + \frac{1}{2\theta^{2n}}((x - y)^2(\theta^2 - 1) + v_n^T S_n^{-1} v_n),$$

with $v_n^{\mathrm{T}} = (-xq, \ldots, y\theta^n(\theta^2 - 1))$. It is easy to see that

$$v_n^{\mathrm{T}} S_n^{-1} v_n = \frac{1}{s_n}(x^2 q^2 s_{n-1} + y^2 \theta^{2n}(\theta^2 - 1)t_{n-1} + 2xy(-1)^n \theta^n(\theta^2 - 1)q^n). \tag{5.24}$$

In addition, it follows from (5.21) and (5.22) that, for any t in the interior of $\mathscr{S}$,

$$\log s_n = n \log a + \log\left(\frac{a - 1}{a - b}\right) + \log(1 - \eta\rho^n),$$

$$\frac{s_{n-1}}{s_n} = \frac{1}{a} - \left(\frac{a - b}{ab}\right)\frac{\eta\rho^n}{1 - \eta\rho^n},$$

$$\frac{t_{n-1}}{s_n} = \frac{1}{a - 1} - \left(\frac{a - b}{(a - 1)(b - 1)}\right)\frac{\eta\rho^n}{1 - \eta\rho^n},$$

with $\rho = b/a$ and $\eta = (b - 1)/(a - 1)$. Finally, (4.3) follows from the conjunction of (5.23) and (5.24) together with the three above expansions, which completes the proof of Lemmas 4.1 and 4.2.

Acknowledgements

References

Anderson, T.W. (1959) On the asymptotic distributions of estimates of parameters of stochastic difference equations. *Ann. Math. Statist.*, **30**, 676–687.

Bahadur, R. and Ranga Rao, R. (1960) On deviations of the sample mean. *Ann. Math. Statist.*, **31**, 1015–1027.

Bercu, B., Gamboa, F. and Rouault, A. (1997) Large deviations for quadratic forms of stationary Gaussian processes. *Stochastic Process. Appl.*, **71**, 75–90.

Bercu, B., Gamboa, F. and Lavielle, M. (2000) Sharp large deviations for Gaussian quadratic forms with applications. *ESAIM Probab. Statist.*, **4**, 1–24.

Cramér, H. (1970) *Random Variables and Probability Distributions*. London: Cambridge University Press.

Dembo, A. and Zeitouni, O. (1998) *Large Deviations Techniques and Applications*, 2nd edition. New York: Springer-Verlag.

Esseen, C. (1945) Fourier analysis of distribution functions. *Acta Math.*, **77**, 1–125.

Grenander, U. and Szegö, G. (1958) *Toeplitz Forms and Their Applications.* Berkeley: University of California Press.

Jensen, J.L. (1995) *Saddlepoint Approximations*. Oxford: Clarendon Press.

Mann, H.B. and Wald, A. (1943) On the statistical treatment of linear stochastic difference equations. *Econometrica*, **11**, 173–200.

Rao, M.M. (1961) Consistency and limit distribution of estimators of parameters in explosive stochastic difference equations. *Ann. Math. Statist.*, **32**, 195–218.

Touati, A. (1996) Vitesse de convergence en loi de l'estimateur des moindres carrés d'un modèle autoregressif (cas mixte). *Ann. Inst. H. Poincaré Probab. Statist.*, **32**, 211–230.

White, J.S. (1958) The limit distribution of the serial correlation in the explosive case. *Ann. Math. Statist.*, **29**, 1188–1197.